ABSTRACT OF THE DISCLOSURE

[0067] An apparatus and method for identifying clusters in two-dimensional data by generating a two-dimensional histogram characterized by a grid of bins, determining a density estimate based on the bins, and identifying at least one cluster in the data. A smoothed density estimate is generated using a Gaussian kernel estimator algorithm. Clusters are identified by locating peaks and valleys in the density estimate (e.g., by comparing slope of adjacent bins). Boundaries (e.g., polygons) around clusters are identified using bins after bins are identified as being associated with a cluster. Boundaries can be simplified (e.g., by reducing the number of vertices in a polygon) to facilitate data manipulation.